DOCUMENT RESUME

ED 117 596 CG 010 298

AUTHOR McGinley, Hugh
TITLE Demand Characteristics in Classical Verbal Conditioning and Attitude Conditioning Studies.
PUB DATE 75
NOTE 48p.; Paper presented at the Annual Meeting of the American Psychological Association (83rd, Chicago, Illinois, August 30-September 2, 1975)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage
DESCRIPTORS *Attitudes; *Classical Conditioning; *Contingency Management; Experimental Psychology; *Learning Processes; Paired Associate Learning; Patterned Responses; Research Projects; *Verbal Learning

ABSTRACT

This paper is a draft for the American Psychological Association Symposium on the conditioning of verbal behavior and attitudes. The author presents the results of several studies he conducted in the classical conditioning of meaning and attitude. These studies attempt to control the measurement effects created by extraneous variables operating on the generalized Staats procedure. Especially at issue is the elimination of the effects of contingency awareness and of demand characteristics in "conditioning of meaning" studies. A problem remaining is determining if contingency awareness is a necessary condition not only for the special case of classical conditioning of meaning but for classical conditioning in general. The author presents a tentative conclusion that the results from Staats' "conditioning of meaning" procedure are not due entirely to demand characteristic effects. (SJL)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED117596

Symposium Draft

Demand Characteristics in Classical Verbal Conditioning and Attitude Conditioning Studies

Hugh McGinley
University of Wyoming

Symposium: Conditioning of Verbal Behavior and Attitudes: A Controversy

American Psychological Association
August 30 - September 3, 1975
Chicago, Illinois

CG 010 258

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

My work in the "classical conditioning" of meaning and attitude started in 1970 and has continued to the present. In all but one of the studies on which I will report, my colleagues and I used an experimental procedure that parallels the generalized Staats procedure. In our procedure we present CVCs (consonant-vowel-consonants) and words or other CVCs in a paired-associate learning task. Unlike the Staats procedure, our procedure calls for repeated associations between conditioned stimuli (CVCs) and unconditioned stimuli (words or CVCs). It is noteworthy that Staats avoided repeated associations in that he believed they would lead to direct associations and possible stimulus substitution--for example, when referring to verbal stimuli, a conditioned stimulus would take on the denotative meaning of an unconditioned stimulus. A resulting measurement of the conditioned stimulus' connotative meaning (usually evaluative meaning a la Osgood, Suci and Tannenbaum, 1957) would be based on the connotative meaning of the unconditioned stimulus. Staats' position is that the connotative meaning of a conditioned stimulus that is established in his experimental procedure is the result of a common meaning component of a variety of unconditioned stimuli that have been associated with a conditioned stimulus and if a specific conditioned stimulus-unconditioned stimulus connection is strong then the conditioned stimulus may simply become a

sign of that unconditioned stimulus. Although we consider Staats' theoretical argument to be logical we believe that it may be overcautious. We decided to use a paired-associate procedure where a conditioned stimulus is associated with a number of unconditioned stimuli of different general meaning but of specific evaluative meaning. Our belief was that, as Staats has theorized, the common evaluative meaning component of the various unconditioned stimuli would be, because of frequency of occurrence, the strongest meaning element associated with the conditioned stimulus and it would be this meaning response rather than the direct recall of a particular unconditioned stimulus that would mediate the evaluative ratings of the conditioned stimulus. Also, we thought that if direct substitution between a conditioned stimulus and a specific unconditioned stimulus were to occur it would be detectable in a postexperimental assessment of awareness.

The first study on which I will report (McGinley & Layton, 1973) used a paired associate list where conditioned stimulus CVCs were paired with either other CVCs or words. Two paired-associate (PA) lists were developed that consisted of four conditioned stimulus CVCs which were each paired with six unconditioned stimuli, either words or CVCs. One of the conditioned stimulus CVCs was paired with CVCs which had been previously rated by university students as having slightly positive evaluative

meanings. A second conditioned stimulus CVC was paired with CVCs that had been rated as having slightly negative evaluative meaning. The remaining two conditioned stimulus CVCs were each paired with three words of slightly positive evaluative meaning and three words of slightly negative evaluative meaning, the overall pairings were essentially algebraically neutral in evaluative meaning. The 12 CVC-word pairs were used as filler items in order to make the PA list more difficult and were not included in the data analyses.

The paired-associates were shown by projecting them on a screen to groups of from three to eight subjects. A conditioned stimulus CVC was shown for 2.3 seconds and then there was a 2.3 second presentation of the conditioned stimulus CVC paired with an unconditioned stimulus, either another CVC or, in the case of the filler items, a word. The interslide interval was approximately .7 of a second. In all, each of the two PA lists consisted of 48 slides (24 associates). The only difference between the two PA lists was that in one of the lists one of the critical conditioned stimulus CVCs was paired with CVCs of positive evaluative meaning while the other conditioned stimulus CVC was paired with CVCs of negative evaluative meaning. In the second list the critical conditioned stimulus CVCs were interchanged. The 35-mm slides were projected by a model 800, Kodak Carousel projector. The

slide exposure time was controlled by a Hunter timer.

The subjects were told that they were involved in a passive learning study where pairs of verbal stimuli, nonsense syllables and words, would be projected on a screen for a short period. They were asked to remember as many of the pairs as they could because they would be asked to recall the pairs at a later time. The subjects were divided into three experimental groups of 21 each. One group of subjects viewed one presentation of the PA list, the second group saw three presentations of the PA list, and the third group viewed five presentations of the list. After the presentation of the PA list, the experimenter asked the subjects to rate 16 nonsense syllables (CVCs) on four 7-interval bipolar scales of evaluative meaning. The scales were pleasant-unpleasant, cruel-kind, beautiful-ugly, and dirty-clean. The scales were combined into a 17-page booklet made up of an instruction page for filling out bipolar scales and 16 pages of the four scales. A different CVC was printed at the top of each page. The two conditioned stimulus CVCs that had been paired with CVCs of either positive or negative evaluative meaning always appeared somewhere between the fourth and eighth pages of the booklet. Within this limitation, the order of the pages of the booklet was random.

After the subjects had completed the ratings they were asked to recall as many of the paired-associates as

they could. Following the recall, the experimenter asked the subjects to fill out a short questionnaire about the experiment. The questionnaire items asked the subjects what procedure or procedures they had used to learn the paired-associates, when they were first aware of using this or these procedures, and what ideas they had about the purpose of the experiment which differed from what the experimenter had stated the purpose of the experiment was (a study about how people learn verbal material). The experimental procedure used no verbal deception, no subject was told anything that, in fact, was not true.

In this study, in order for subjects to be contingency aware they would not only have to be aware of the absolute pairings but they would also have to be aware that unconditioned stimulus CVCs which had been paired with critical conditioned stimulus CVCs had common positive or negative evaluative meaning. Subjects would have had to be contingency aware before they could have been demand aware, i.e., aware of the relationship between the pairing of the conditioned stimulus CVCs with other CVCs of either positive or negative evaluative meaning and their subsequent evaluative ratings of the conditioned stimulus CVCs on the bipolar scales.

The conditioning data (for convenience, we will refer to CVC ratings on the bipolar scales as a measure of conditioning) were collected in two forms, the conditioned

stimulus CVC ratings from the first of the four bipolar scales (pleasant-unpleasant) and mean ratings from the total of the four bipolar scales. Analyses of the pleasant-unpleasant scale data showed that conditioning occurred only for subjects who saw five trials of the PA list. The data from the total of the four bipolar scales showed the same result. The mean pleasant-unpleasant scale ratings are presented in Table 1 and Figure 1. Of the 12 possible critical CVC pairs, subjects who viewed one trial recalled an average of .33 pairs, subjects who viewed three trials recalled an average of 1.23 pairs, and subjects who viewed five trials recalled an average of 3.95 pairs. The subjects' responses to the questionnaire indicated neither contingency nor demand awareness. Most subjects said that they tried to learn the pairs either by rote memory or by using phonic similarities of the pairs.

Table 1

Mean pleasant-unpleasant ratings of the critical conditioned stimulus CVCs

	Trials[a]		
Ratings	1	3	5
CVC-CVC (positive)	4.05[b]	4.14	4.62
CVC-CVC (negative)	3.95	4.23	2.86
difference	.10	- .09	1.76**

[a]There were 21 subjects in each trial condition.
[b]Based on a 7-point scale, high numbers indicate positive ratings.
**$p < .01$

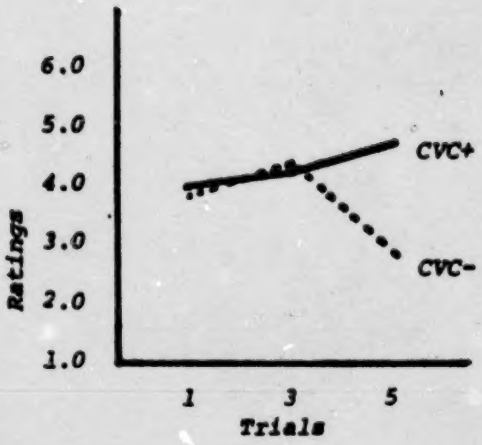

Figure 1
Mean pleasant-unpleasant ratings of critical conditioned stimulus CVCs

McGinley and Layton (1970) conducted a study that was identical in design to their CVC-CVC study other than the critical pairings were CVC-word pairs while the filler items were CVC-CVC pairs. One conditioned stimulus CVC was consistently paired with words of evaluatively positive meaning while another conditioned stimulus CVC was consistently paired with words of evaluatively negative meaning. The two filler conditioned stimulus CVCs were paired with CVCs that were neutral in evaluative meaning. On the basis of the subjects' responses to the awareness questionnaire, judges rated 11 of the 61 subjects (18%) as contingency aware. Of these, one subject was in the 1-trial condition, three subjects were in the

3-trial condition, and seven subjects were from the 5-trial condition. No subject was judged as demand aware. Analyses of the pleasant-unpleasant scale data showed a conditioning of meaning effect for all groups of subjects. The combined ratings from the four bipolar scales yielded similar results. The average pleasant-unpleasant ratings are shown in Table 2 and Figure 2.

Table 2

Mean pleasant-unpleasant ratings of the critical conditioned stimulus CVCs

	Trials		
Ratings	1(18)[a]	3(17)	5(15)
CVC-word (positive)	5.00[b]	5.00	5.27
CVC-word (negative)	3.50	2.41	2.20
difference	1.50**	2.69**	3.07**

[a] The number of subjects in each trial condition is in parenthesis. All of these subjects were classified as not being contingency aware.

[b] Based on a 7-point scale, high numbers indicate positive ratings.

**$p < .01$

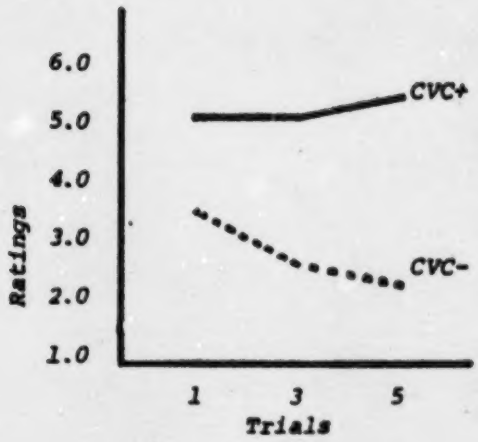

Figure 2
Mean pleasant-unpleasant ratings of critical conditioned stimulus CVCs

After the second study we chose to manipulate the difficulty of the PA list by decreasing the slide presentation time. We designed a study (McGinley, Layton & McGinley, 1971) which was similar in mechanical design to the two previous studies except that the stimuli presentation time was 1.7 seconds and there were only four conditioned stimulus CVCs. One of the critical conditioned stimulus CVCs was paired with six evaluatively positive words and an evaluatively neutral CVC while the other critical conditioned stimulus CVC was paired with six evaluatively negative words and an evaluatively neutral CVC. The remaining two CVCs were each paired with three evaluatively slightly positive words, three

evaluatively slightly negative words and an evaluatively neutral CVC. In all, the list consisted of 28 paired associates. The two critical conditioned stimulus CVCs were counterbalanced across the evaluative meaning words so that a possible order effect could be assessed. Control group subjects rated the CVCs on the bipolar scales before they viewed the PA list. With counterbalancing, there were eight groups of 12 subjects each. Eight of the 72 experimental group subjects were judged as contingency aware on the basis of their responses to the awareness questionnaire. Data from these subjects were not included in the analyses. The analyses of both the pleasant-unpleasant ratings and the total scale ratings showed a conditioning effect and a trials effect. The mean pleasant-unpleasant ratings for each of the critical conditioned stimulus CVCs are shown in Table 3 and Figure 3. There was no effect for order.

Table 3

Mean pleasant-unpleasant ratings of the critical conditioned stimulus CVCs

	Trials			
Ratings	control (24)[a]	3 (20)	5 (22)	9 (22)
CVC-word (positive)	4.17[b]	4.92	5.55	5.77
CVC-word (negative)	4.21	4.08	2.77	2.91
Difference	- .04	.84*	2.78**	2.86**

[a] The number of subjects in each group is in parenthesis. All of these subjects were classified as not being contingency aware.

[b] Based on a 7-point scale, high numbers indicate positive ratings.

* $p < .05$

** $p < .01$

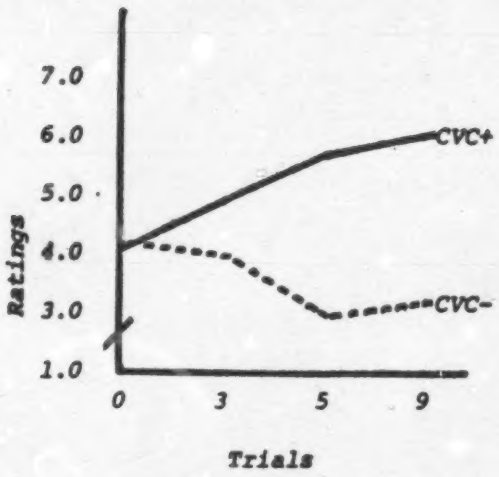

Figure 3
Mean pleasant-unpleasant ratings of critical conditioned stimulus CVCs

In none of the three studies presented thus far did we judge subjects as demand aware. In the next study (McGinley & Lloyd, 1971) we manipulated the type of awareness questionnaire used. There were three experimental groups and a control group. The PA list was an easier one than those used previously. It consisted of 18 pairs. One CVC was paired with six evaluatively positive words, another with six evaluatively negative words and a third with three evaluatively slightly positive words and three evaluatively slightly negative words. The slide presentation time was 2.3 seconds. Because we found no order effect in a previous study (McGinley, et al., 1971) we did not counterbalance condition stimulus

CVCs across evaluative meaning words. All subjects saw five presentations of the PA list. The control group subjects rated the CVCs before they saw the PA list. The questions from the three awareness questionnaires are given in Table 4.

Table 4

Awareness Questionnaires

Awareness Questionnaire I (AQ-I)

1a. What procedure (or procedures) did you use to learn the pairs?
b. At what time during the experiment were you aware that you were using the learning procedure you described above?
2. In the past we have found that some people who participate in verbal learning experiments form ideas of hypotheses about how or what they should learn, etc., and these ideas are independent of the experiment's instructions. Did you form such ideas? Yes _____ No _____ (check one)
3. If you checked yes to question #2, what was your hypothesis (or hypotheses)?
4. When did you develop the hypothesis? Check one of the following:
_____ During the presentation of the slides.
_____ After the presentation of the slides.
_____ While reading this questionnaire.

Awareness Questionnaire II (AQ-II)

1. What procedure (or procedures) did you use to learn the pairs.
2. At what time during the experiment were you aware that you were using the learning procedure you described in the above question?
3. The Experimenter is obviously interested in the number of pairs that you can recall. What else do you think he may be interested in?
4. Please answer the following questions as truthfully as possible--we are interested in WHAT YOU THINK!
a. Why were nonsense syllables paired both with meaningful words and words with little meaning?

b. Why did subjects meet in small groups?
c. Did being with other people affect your learning performance?
d. Why were you asked to fill out the rating scales immediately after you saw the list?
e. Were the rating scales difficult for you to fill out?

Awareness Questionnaire III (AQ-III)

This questionnaire was the same as AQ-II except that item (f) was added to question number four.

f. Did you fill out the rating scales as to how you felt about the nonsense syllables or did you fill them out as to how you thought the experimenter wanted the rating scales to be filled out? If the latter, what was your thought on how the experimenter wanted the scales to be filled out?

Based on the subjects' written responses to the awareness questionnaires the following percentages of subjects were judged as contingency aware: AQ-I, 38%; AQ-II, 48%; AQ-III, 56%. We found the demand aware classification harder to make than the classification of contingency aware and consequently decided on the trichotomy of aware, maybe-aware and unaware. We used the maybe-aware classification for subjects who wrote something about the CVC ratings but who did not relate the CVCs to the words with which they were paired. None of the subjects who filled out AQ-I were judged as demand aware and only two of them were judged as maybe-aware while the percentages of subjects aware or maybe-aware from the other two groups were: AQ-II, 32% aware, 40% maybe-aware; AQ-III, 56% aware, 32% maybe-aware. Analyses of the pleasant-unpleasant ratings of the CVCs showed that subjects who were classified as unaware (contingency or

demand) showed conditioning. Of course, those subjects who were classified as aware also showed conditioning. The results of the analyses of the total scales data were similar to the single scale results. Table 5 lists the mean pleasant-unpleasant scale ratings for subjects judged for contingency awareness while Table 6 lists the mean ratings for subjects judged for demand awareness. These two sets of data results are shown in Figures 4 and 5 respectively.

Table 5

Mean pleasant-unpleasant ratings from subjects judged for contingency awareness

Questionnaire	CVC-word (positive) aware	CVC-word (positive) unaware	CVC-word (negative) aware	CVC-word (negative) unaware
AQ-I(24)[a]	6.67(9)[bc]	6.07(15)	1.59(9)	2.40(15)
AQ-II(25)	5.83(12)	5.77(13)	2.50(12)	2.15(13)
AQ-III(25)	5.93(14)	5.73(11)	1.64(14)	2.73(11)
control(26)	4.08	4.08	3.70	3.70

[a] The number of subjects in each group is in parenthesis. e.g., there were 24 subjects in AQ-I, 9 of whom were judged to be contingency aware.

[b] Based on a 7-point scale, high numbers indicate positive ratings.

[c] As compared to the control group, all experimental groups, regardless of awareness classification, showed conditioning of evaluative meaning.

Table 6

Mean pleasant-unpleasant ratings from subjects judged for demand awareness

Questionnaire[a]	CVC-word (positive)			CVC-word (negative)		
	aware	maybe-aware	unaware	aware	maybe-aware	unaware
AQ-II & AQ-III (50)[b]	6.18(22)[c]	5.33(18)[d]	5.90(10)	2.00(22)	2.39(18)	2.40(10)
control(26)	4.08	4.08	4.08	3.70	3.70	3.70

[a] Subjects from AQ-I were not included as AQ-I did not lead to demand aware classifications of the subjects.
[b] The number of subjects in each group is in parenthesis.
[c] Based on a 7-point scale, high numbers indicate positive ratings.
[d] As compared to the control group, all groups showed conditioning of evaluative meaning.

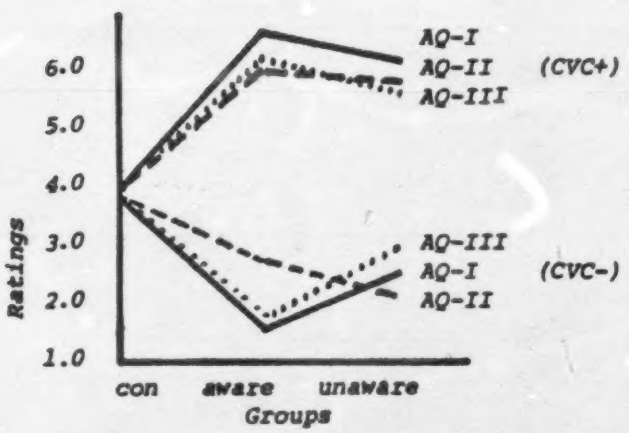

Figure 4
Mean pleasant-unpleasant ratings from subjects judged for contingency awareness

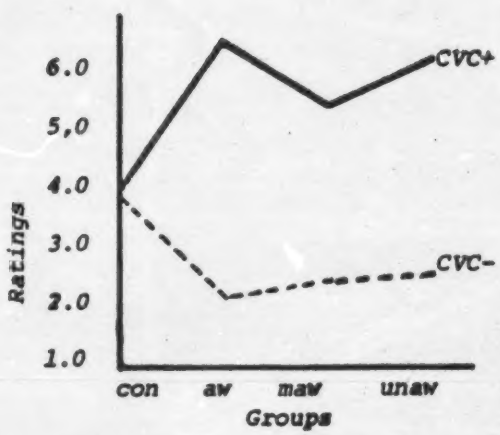

Figure 5
Mean pleasant-unpleasant ratings from subjects judged for demand awareness

McGinley and Boone (1974) conducted a conditioning of meaning study where the dependent measure, ratings of CVCs that had been associated with words of either evaluative positive or negative meaning, was separated from the conditioning stage of the experiment. I will go into detail about the experimental procedure of the study as it varies considerably from the other studies on which I have reported.

A 24-item PA list was developed where two CVCs were each paired with: (1) evaluatively positive words (the nouns music, grace, lake and freedom with one CVC; and valley, bath, butterfly and deer with the second CVC), (2) evaluatively negative words (the nouns rape, infection, fraud and jealousy with a third CVC; and flea, sword, anarchy and delirium with a fourth CVC), and (3) evaluatively neutral words (the nouns unit, code, hint and iron with a fifth CVC; and train, ink, box and ounce with a sixth CVC). As before, the list was presented with an anticipation of presentation procedure where a slide of a conditioned stimulus CVC was followed by a slide of the CVC paired with a noun. Each slide was presented for 2.5 seconds, the interslide time was approximately .75 of a second. The 24-item PA list was presented five times. The presentation of the PA list was videotaped and all experimental subjects viewed the playback of the videotape on television sets with 23-inch screens.

Subjects were solicited frcm five introductory psychology classes to participate in a learning experiment. Subjects from three of the classes met on three occasions. During the first meeting they saw and heard a videotape playback of their experimenter (Boone) who gave instructions for the experiment. The subjects then saw the presentation of the PA list. Following this they recalled the items of the PA list. Two days later they met again. They first recalled as many of the CVC-word pairs as they could then they watched the television presentation of the PA list and then they recalled the pairs again. The next day, during a meeting of their introductory psychology classes, the subjects participated in the second part of a project called project NORMS. Approximately one month earlier students in all introductory psychology classes had participated in a word rating project where words were rated on a series of bipolar scales. The NORMS experimenter (McGinley) reappeared and said that he was also developing norms for a number of nonsense syllables and he then asked the students to complete the rating forms (students who did not wish to do the ratings left the classroom for the remainder of the class meeting time). The students rated 20 nonsense syllables on nine, 9-interval bipolar scales (six of the scales had evaluative meaning bipolars). Five days later the subjects met for a third time. During this meeting

they rated 12 CVCs on four, 7-interval bipolar scales of evaluative meaning, recalled the CVC-word pairs and filled out an awareness questionnaire.

Subjects from a fourth introductory psychology class met twice. These subjects experienced the same experimental procedure up to the end of the second learning session. At that time they rated the CVCs and filled out an awareness questionnaire. These subjects did not participate in the second session of project NORMS. Volunteers from a fifth introductory psychology class served as controls. They met only once. They first rated the CVCs and then viewed the PA list. After the presentation of the list they recalled the pairs.

As might be expected, attrition was high. Analyses were based on data from 40 subjects for whom we had both PA learning data and NORMS data, on 16 subjects who attended two meetings of the learning experiment but for whom NORMS data were not collected, and on 35 subjects in the experimental control group. A NORMS-rating control group was obtained by using data from 40 subjects who did not participate in the learning experiment. There were two general kinds of data analyses, one for CVC ratings made by subjects during the learning experiment and another for CVC ratings made during project NORMS. In both cases we analyzed ratings from the pleasant-unpleasant scale and from a total of four evaluative meaning scales. These

data, presented as group averages, are given in Tables 7, 8, 9 and 10, and in Figures 6, 7, 8, and 9.

Table 7

Mean pleasant-unpleasant ratings at the end of the learning experiment

Groups	CVC-word (positive)[a]	CVC-word (negative)	CVC-word (positive) (negative)
Experimental (56)[b]	5.55[c]	3.71	4.10
Control (35)	4.51	4.06	4.16
difference	1.04**	- .35*	.06

[a] The data for the two CVCs in each reinforcement condition are combined.
[b] The number of subjects in each group is in parenthesis.
[c] Based on a 7-point scale, high numbers indicate positive ratings.
* $p > .05 < .10$
** $p < .01$

Table 8

Mean total evaluative scale ratings at the end of the learning experiment

Groups	CVC-word (positive)[a]	CVC-word (negative)	CVC-word (positive) (negative)
Experimental (56)[b]	21.46[c]	13.99	16.11
Control (35)	17.80	16.20	16.59
difference	3.66**	- 2.21**	.48

[a] The data for the two CVCs in each reinforcement condition are combined.
[b] The number of subjects in each group is in parenthesis.
[c] Based on a total of 28 points, high numbers indicate positive ratings.
** $p < .01$

Table 9

Mean pleasant-unpleasant ratings from project NORMS

Groups	CVC-word (positive)[a]	CVC-word (negative)	CVC-word (positive) (negative)
Experimental (40)[b]	6.64[c]	4.15	4.93
Control (40)	5.51	5.28	4.86
difference	1.13**	-1.13**	.07

[a]The data for the two CVCs in each reinforcement condition are combined.
[b]The number of subjects in each group is in parenthesis.
[c]Based on a 7-point scale, high numbers indicate positive ratings.
**$p < .01$

Table 10

Mean total evaluative scale ratings from project NORMS

Groups	CVC-word (positive)[a]	CVC-word (negative)	CVC-word (positive) (negative)
Experimental (40)[b]	25.99[c]	17.24	18.81
Control (40)	21.10	20.43	19.54
difference	4.89**	3.19**	.73

[a]The data for the two CVCs in each reinforcement condition are combined.
[b]The number of subjects in each group is in parenthesis.
[c]Based on a total of 36 points, high numbers indicate positive ratings.
**$p < .01$

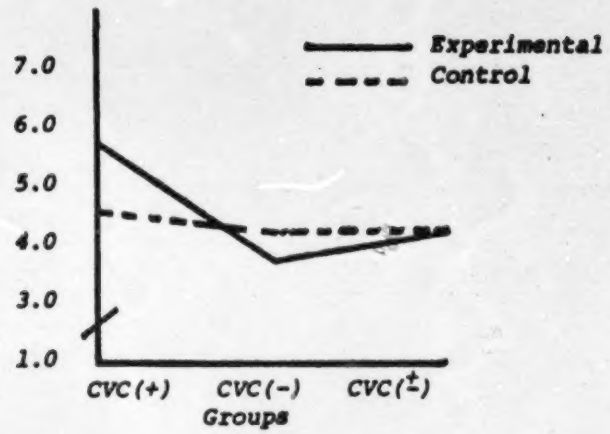

Figure 6
Mean pleasant-unpleasant ratings at the end of the learning experiment

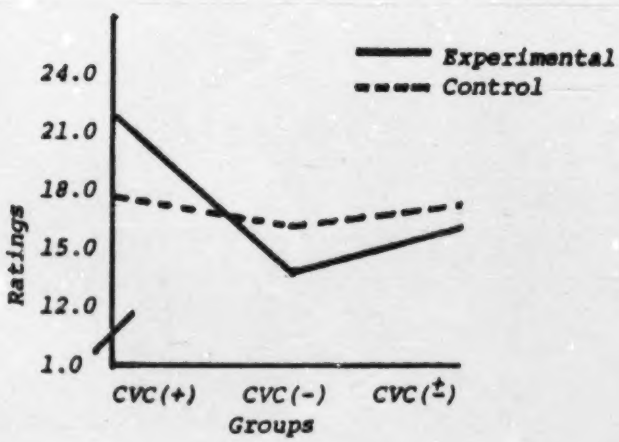

Figure 7
Mean total evaluative scale ratings at the end of the learning experiment

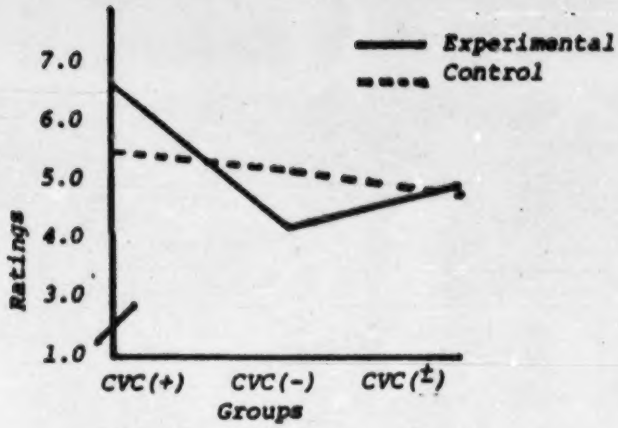

Figure 8
Mean pleasant-unpleasant ratings from project NORMS

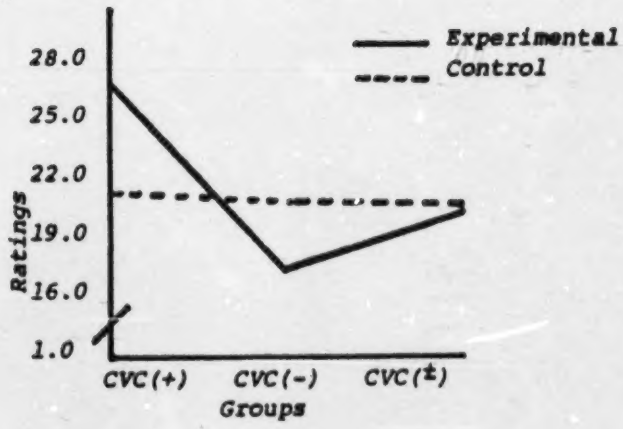

Figure 9
Mean total evaluative scale ratings from project NORMS

As would be expected from the results of the previous studies, the results of the present study show a conditioning of meaning effect when the CVC ratings were collected during the last meeting of the learning experiment. However, there was also a conditioning of meaning effect when the ratings were collected outside of the experimental setting, in the classroom while the subjects were supposedly participating in a norm gathering project.

There is another study that I wish to present in this series. However, since it is an experiment which differs from the four studies that I have presented so far, it is appropriate to discuss the previous studies before proceeding to it.

There are various strategies that may be taken when trying to eliminate the effects of contingency awareness and of demand characteristics in conditioning of meaning studies. The likelihood of detecting artifactual effects that may be attributed to demand characteristics is within our technical knowhow. Our problem here is simply one of thoughtful experimentation including a careful consideration of controlling measurement effects created by extraneous variables. The testy problem, to me at least, is to determine if contingency awareness is a necessary condition for not only the special case of classical conditioning of meaning

but for classical conditioning per se. Indeed, Brewer (1974), after an extensive review of the classical conditioning literature, believes that there is no hard evidence for classical conditioning (in adult humans) without contingency awareness. It may be that a reassessment of the role of contingency awareness in classical conditioning studies is in order. However, a reassessment may lead to the conclusion that we have had very few studies which test either the hypothesis that contingency awareness is not necessary for classical conditioning in adult humans or the alternative hypothesis that it is necessary. Certainly, a person's inability to verbalize an association between two or more stimuli at the time of their occurrence plus a lag time does not mean that the person could not have verbalized the association at its occurrence. In the same vein, a person who verbalizes an association between stimuli at some time after its occurrence does not mean that the person could have verbalized the association at the time of its occurrence.

I would like to pose that we have at least two categories of contingency awareness and that each of these categories has at least two stages. Stage one of the first category would include awareness of an association between two or more stimuli at the time of their occurrence. This awareness, let's call it

contingency awareness-Ia (CA-Ia), would be the central issue of consideration in assessing the necessity for contingency awareness in learning. The second stage of contingency awareness-I would be the awareness of an association between stimuli at a time after the stimuli occurred (CA-Ib). Contingency awareness-IIa would be an awareness of properties of associated stimuli at the time of their occurrence while contingency awareness-IIb would be such an awareness at a time after the occurrence of the stimuli.

Let us use the classical conditioning of meaning procedure to clarify these four incidences of contingency awareness. A conditioned stimulus CVC is temporally paired with another stimulus, a word (or object, emotion, etc.). CA-Ia would be that the person is aware of both the CVC and the word at the time of their occurrence (and association). CA-Ib would be that the person demonstrates awareness of the stimuli at a later time. CA-IIa would be that a person, at the time of the occurrence of the CVC and word stimuli, is aware of properties of the word stimulus, say connotative meanings, and that these properties are associated with the CVC. CA-IIb would be an awareness of the properties of the word and their association with the CVC at a later time. With these categories of contingency awareness in mind, let us consider demand awareness.

Demand characteristics can be either intrinsic to or extrinsic of an experiment. In the classical conditioning of meaning and attitude studies investigators have usually attempted to deal with intrinsic demand characteristics. Extrinsic demands refer to possible influences created by actual or fictitious information about what is expected in a particular experiment, misconceptions about what psychologists study (apprehension of evaluation) and general cultural or ethnic group characteristics. Intrinsic demands relate to influences on subjects' behaviors by characteristics of the formal and informal procedure of an investigation. Rosenthal's (1966) work in experimenter expectancy effects is an example of informal procedural demands while Staats' experimental procedure for the conditioning of meaning and attitudes (my procedure too!) and Page's assessment of demand awareness procedures (mine too!) have facets that are examples of formal procedural demands that may lead to artifactual "experimental" effects.

It seems to me that, in the typical classical conditioning of meaning and attitude study, an awareness of intrinsic demand characteristics would presuppose either CA-Ib or CA-IIb or both. In the case of CA-Ib, a subject may view a CVC as a sign for a known word and then treat the sign as the word (a sign developed for another sign, the word). This may lead to the belief that the CVC has

taken on the properties of the word. We should not classify this sort of CVC-for-a-word substituttion as concept formation in that the CVC means no more or no less than the word for which it stands. Staats, in his conditioning procedure, has attempted to reduce sign substitution by creating weak associations between conditioned stimuli and various unconditioned stimulus words of different denotative meaning while, at the same time, creating a strong association between conditioned stimuli and a common meaning component of the various words, usually evaluative meaning.

If there were no CA-Ib (regardless of CA-Ia) and a subject rates a CVC as if it were the word with which it had been paired previously, then, through conditioning (or substitution) connotative meaning has been established for the CVC. Yavuz and Bousfield (1959) found that connotative meaning remained associated with paralogs (supposedly, Turkish words) even though subjects could not recall the actual words with which the paralogs had been paired. Yavuz and Bousfield's subjects were CA-Ia when they learned the paralog-word pairs but not CA-Ib when they rated the paralogs on evaluative meaning scales.

CA-IIa, like CA-Ia, has not been isolated well and there is not much we can say about it. Our main concern is CA-IIb since most awareness assessments of common connotative meaning amongst various words (unconditioned

stimulus words) are made after a pairing manipulation. CA-IIb is the precursor of demand awareness that is assessed postexperimentally. Subjects who verbalize a relationship between evaluative meaning ratings scales and ratings of a particular CVC are probably aware, at least at the time of their verbalizations, of the CVC's association with words possessing common meaning components. However, the subjects actually may not have been aware of such a relationship when they rated the CVC, that is, something may have occurred after the rating that lead to a verbalization of CA-IIb (or of CA-IIa). Also, just to complicate matters, the subjects could have been CA-IIb while rating the CVC but not CA-IIb at the time their awareness was assessed. Furthermore, persons who are CA-IIb and demand aware may not have rated a CVC in a particular way "because they were suppose to do so" but because they "felt that way about the CVC." Conditioning of meaning effects may be the result of variables other than demand awareness even though subjects are able to verbalize an awareness of demand characteristics. At this point, the waters of great knowledge are both troubled and muddied.

A common strategy in conditioning of meaning studies has been to assume that subjects who are considered to be CA-IIb, and sometimes CA-Ib, could be responding to rating scales on a basis other than conditioning. Data obtained from these subjects are either not analyzed or are analyzed

separately. Such a procedure would be experimentally sound if we knew that our judgments of CA-Ib and CA-IIb were valid. We do not know about the validity of our judgments but, nevertheless, we must proceed about our business as psychologists and accept the fact the we err some of the time. Sometimes we will call subjects CA-IIb and/or CA-Ib when they are not (false positive) while at other times we will classify subjects as not being CA-IIb and/or CA-Ib when indeed they are (false negative). There is no way of knowing for sure but I suspect that judges who identify with behavior theory are guilty of obtaining more false negatives than they should while judges who follow cognitive theory obtain more false positives than they should. Both types of errors are serious and should be reduced whenever possible.

My training has been more in line with behavior theory than with cognitive theory. In keeping with this, I suspect that my research is somewhat behavioristic in character and may be insensitive to some issues that are extremely important to the more cognitively oriented psychologists. Whatever, I will now offer my interpretations of the results of studies that I have presented.

The CVC-CVC study (McGinley & Layton, 1973) was an early study that was conceived of and conducted while I was searching for a workable alternative to the Staats procedure for the conditioning of meaning. I have not

attempted to replicate the study. Jim James (1973, personal correspondence) did a replication of sorts and got ambiguous results. He, however, conducted the experiment using a large number of subjects (40-50) in a single setting, a procedure which, I feel, results in low subject motivation (the "pinch me to see if I am alive" syndrome). James also accepted the average evaluative meaning ratings of CVCs that I had gathered from Canadian students--he probably should have checked these ratings for consistency with his south-midwestern student population. The evaluative meaning commonality of the unconditioned stimulus CVCs that were used in the study were very subtle, the subjects responses to the awareness questions indicated a fairly high degree of frustration--the list was very difficult to learn, after five presentations of the PA list the correct paired-associate recall was only 3.95 pairs out of 12 pairs (about 25%). I believe that demand awareness and CA-IIb were reduced by the procedure of this experiment, however, there was a 25% recall and thus, by definition, CA-Ib. It could be that some sort of stimulus substitution is what caused the "conditioning" effect.

The first of the CVC-word studies (McGinley & Layton, 1970) was the sister work of the CVC-CVC study. The CVC-word pairing procedure seemed to have a better potential as an experimental procedure because it was

not, seemingly, quite as frustrating as the CVC-CVC procedure. Our short awareness questionnaire yielded responses from 18% of the subjects that caused us to classify them as either CA-Ib or CA-IIb. Only one subject was classified as demand aware. When the judged aware subjects were dropped from the analyses there was still a conditioning effect for subjects who saw either 1, 3, or 5 trials of the PA list. The brief awareness questionnaire seemed to be completely insensitive to demand awareness (this statement assumes that at least some demand awareness should have been present). Also, the observed conditioning could have been the result of sign substitution if we assume that the postexperimental awareness questionnaire was not sensitive to CA-Ib.

The second CVC-word study (McGinley, Layton & McGinley, 1971) used the general procedure of the previous CVC-word study other than the PA list consisted of four CVCs and their associates instead of six CVCs plus associates. Also, we reduced the slide presentation time from 2.3 seconds to 1.7 seconds. Whatever the phenomenological effect, the judged contingency awareness was 11% and the awareness questionnaire did not elicit or provoke responses that would have been judged as suggesting demand awareness.

In the first three of our experiments we attempted to vary awareness by the complexity of the PA list and by

varying slide time. In the fourth experiment (McGinley & Lloyd, 1971) we studied the conditioning of meaning phenomenon by using more extensive awareness questionnaires in addition to the questionnaire that we had used previously. In that we wanted to set conditions for greater incidences of CA-Ib, CA-IIb and demand awareness we slowed the slide presentation time to 2.3 seconds and used only three conditioned stimulus CVCs, each of which was paired with six words. This procedure led to a much greater number of subjects who were classified as contingency aware. This was true for all of the experimental groups including a group who filled out the questionnaire that we had used in the previous studies--38% of this group received contingency awareness judgments. The briefer questionnaire, however, still did not lead to subjects being classified as demand aware. Forty-eight and 56% of the subjects who responded to the other two awareness questionnaires were classified as contingency aware. Of the 50 subjects who responded to the latter two questionnaires, 44% were judged to be demand aware, 36% maybe-aware and 20% were not thought to be demand aware. All subjects, however, showed conditioning effects. On the surface, these results seem to support the viewpoint that reportable contingency awareness is not needed for the classical conditioning of meaning. On the other hand, one might want to say that all of the subjects were

probably contingency aware but only some of them wrote responses to the awareness questions that would cause the judges to classify them either as CA-Ib or CA-IIb. So goes the difficulty of assessing awareness. It is interesting to note that, in this study, our first-used awareness questionnaire resulted in awareness judgments for 38% of the subjects who responded to it. There seems to be a definite relationship between the difficulty of the PA learning task and reported awareness when the awareness questionnaire is held constant.

The last of the studies that was presented earlier is a study by McGinley and Boone (1974). To me the results of this study are quite clear--the conditioning of meaning results cannot be attributed solely to subjects who are demand aware. The CVC ratings that were obtained in the classroom under the guise of a project for developing rating norms for words and nonsense syllables showed a conditioning or development of meaning effect for subjects who had also participated in a conditioning of meaning study. According to my understanding of demand awareness and its effect, subjects who are CA-Ib or CA-IIb may make a connection between an experimental manipulation and the measure of the effect of that manipulation. Once the connection is made the subjects respond in the manner that they believe the experimenter expects them to respond--CVCs that have

been paired with words that connote positive evaluation are suppose to be rated positively on bipolar scales, etc. In the present study even though there was no extensive connection between the independent variable manipulation and the measure of the possible effect of that manipulation the subjects still rated CVCs in accordance with CVC-word pairings that they had been exposed to in another context. This finding strongly suggests that although previous studies have shown a substantial correlation between judged demand awareness and the development of meaning, the suggested relationship is not one of cause and effect. I need to add, however, that the data from this study do not clarify how meaning development occurs. A conditioning of meaning model as well as a concept-formation model (Rhine, 1958) would equally account for the rating effect found in this study. Also, it is interesting to note that the interpretation that I have made from the results of this study are in accord with the results of a similar study by Zanna, Kiesler and Pilkonis (1970) where there was an attempt to dissociate the dependent measure from the independent variable manipulation stage of a conditioning of meaning study.

The last study of this series on which I wish to report is a recent experiment by Boone and McGinley (1975). The study relates to the conditioning of affect to visual

stimuli of red and blue lights and to properties of the stimuli, their color and words denoting their color. The direct measure of affect was GSR while two indirect measures were preference for red or blue colored stimuli in a problem solving task and evaluative meaning ratings of the words red and blue. An extensive postexperimental questionnaire was used in order to estimate contingency and demand awareness of subjects.

The experiment consisted of three phases, conditioning, assessment of conditioning effects, and assessment of awareness. The experimental procedure is similar to that used by McGinley (1970). Forty male subjects completed a tone discrimination task where incorrect responses were punished by electric shock and correct responses were rewarded with five-cents (a nickle). There were lights on an instrument panel in front of the subject. For half of the subjects, a red light preceded electric shock by .5 of a second and a blue light preceded the delivery of a nickle by .5 of a second. The light-color and reinforcement type were reversed for the other half of the subjects. The conditioning procedure continued until a subject demonstrated GSR conditioning to the light associated with electric shock (during special test trials). In the second phase of the experiment, subjects solved a figure discrimination task where they viewed a series of pictures of unfolded geometric figures

each followed by pictures of five folded figures. The figures were presented for ten seconds during which time the subjects had to choose which of the five folded figures would be a folded version of the unfolded figure. Two of the five folded figures were correct but one was both inversed and reversed from the other. The five folded figures were lightly colored. On critical slides, one of the correct figures was red while the other was blue. We hypothesized that since the task was so rushed, subjects would have to make snap judgments and these judgments would be influenced by affect that would be elicited by red and blue hues, i.e., subjects would tend to choose figures that were of the hues that were associated with either reward or punishment (the approach towards the reward-associated colored figure should have been enhanced by an avoidance of the punishment-associated colored figure). Following this task the subjects filled out an awareness questionnaire (see Table 11). They then rated 12 words on bipolar scales, included in the words were red and blue. Forty male subjects served as a control group. These subjects experienced phases II and III of the experiment. For analyses purposes they were arbitrarily separated into two groups of 20 subjects each.

Table 11

Awareness Questionnaire

(see attached sheets)

The results of the study are not clear but they suggest a conditioning of affect effect. The average number of visual discrimination correct responses for red and blue colored figures are listed in Table 12.

Table 12

Average correct visual discrimination task responses for experimental and control subjects

Reinforcement condition	Blue colored figures		Red colored figures	
	Experimental Ss (n=20)	Control Ss (n=20)	Experimental Ss (n=20)	Control Ss (n=20)
Reward	3.30	2.40	1.90	2.30
Punishment	2.65	2.35	1.90	2.30
difference	.65*	.05	.00	.00

*$p < .05$

The correct responses results are not consistent for figures of the two colors, red and blue. The most critical comparisons to test are those between figures of the same colors. The data for blue figures support the hypothesis while those for red figures do not. Control group data are for reference and not direct comparisons. As is seen in Table 12, control subjects chose red figures as often as they chose blue figures.

The experimental subjects, however, chose red figures less often than blue figures regardless of affect association. Overall, of the two important comparisons, blue (+) vs blue (-) and red(+) vs red (-), only the data of the blue figures support a conditioning of affect hypothesis. Table 13 gives the mean evaluative meaning ratings for the words red and blue.

Table 13

Mean evaluative meaning ratings of red and blue for experimental and control groups

Word	Color associated with reward	Color associated with punishment	Control ratings
Blue	17.35[a][b]	18.85	18.85
Red	24.30	26.55	25.32

[a] The total scale maximum score is 54, there were six 9-point scales.
[b] Low scores indicate positive ratings.

For word ratings, the control subjects rated blue more positively than they rated red. This trend held for the experimental group subjects too. No other set of means differed significantly although both of the affect within color comparisons were in the appropriate direction--the reward associated colors were rated more positively than were the punishment associated colors. Also, when the control group's ratings are used as expected ratings, blue(+) and red(+) are rated more

positively than expected while blue(-) and red(-) are rated less positively than expected. Although no one of these tests for conditioning resulted in a statistically significant difference, six out of six of these tests are in the direction which support a conditioning of affect thesis.

The results of the study offer limited support to the proposition that affect (or components thereof) that is conditioned to a specific stimulus (red or blue light) will also be conditioned to properties of the stimulus, in this study, hue and word-sign.

It seems quite obvious that adult humans are capable both of learning associations between stimuli and of demonstrating that such associations have been learned. This human ability, according to some behavior theorists, need not be related to a person's ability to verbalize what has been learned (associated), that is, in a very loose generalization, some complex learning is "precognitive"if we delimit cognition to language correlated responses.

The behavior theorists seem to be of two camps of thought (if indeed there is enough agreement among research psychologists to create "theoretical camps"). One camp of behavior theorists rarely seeks explanations for directly observable behavior beyond description of consistent events that precede specific behavior. Other

behaviorists are deeply concerned with what happens between external antecedent and consequential variables and evoke directly-unobservable, internal variables for explanatory purposes. Staats' theory of the development of meaning and attitude is seemingly consistent with the latter theoretical camp.

The experimental procedure that Staats used in his conditioning of meaning studies (e.g. 1957, 1958a, 1958b, 1959) clearly indicates that he was aware of demands that can be created in the experimental situation which may lead to artifactual results. For instance he used a ruse about what subjects were suppose to do in the experiment (learn material that was presented in two sensory modalities) and questioned subjects about their awareness of common evaluative meaning amongst stimuli.

Most criticisms that have been voiced by researchers such as Page (1975) have been directed toward Staats' experimental procedure and his interpretation of data rather than toward Staats' theory of meaning and attitude development and change per se. Page believes that Staats' experimental procedure contains strong demand characteristics and that he has ineffectively attempted to control for the possible effects of these demands. I, personally, understand Page as saying that Staats' experimental procedure cannot effectively test his (Staats') theory of meaning and attitude development. I would be quite

surprised if Page were to state that Staats' procedure leads to results which support a cognitive theory of the development of meaning and attitude. Other investigators, however, have suggested that the results of some conditioning of meaning studies support a cognitively oriented theory of concept formation (Gerstein, 1961; Hare, 1964, 1965; Insko & Oakes, 1966; O'Donnell & Brown, 1973; O'Donnell, 1975).

Conclusions are rather difficult to draw as they are merely labels that are attached to many sets of data much in the same way that we label factors from factor analyses. Whatever the contrivance, we must make at least tentative conclusions. Although I do not believe that, at this time, I am in a position to make a statement about a possible concept formation explanation of the "conditioning" results commonly obtained from Staats' conditioning of meaning procedure, I will walk the line and say that the conditioning results are not entirely due to demand characteristic effects. I base this tentative conclusion on my own work and studies by O'Donnell (1975), and Zanna et al. (1970).

Stage III Questionnaire

SET A: Experimental Subjects

1. Explain briefly what you think the tone discrimination experiment was about and how you think we wanted you to reach the goal of the experiment.

2. In the tone discrimination experiment how sure were you on your correct choices?

a. Absolutely sure
b. Pretty sure
c. Better than 50% sure
d. 50/50 sure
e. Guessing, but kind of sure
f. Guessing
g. Blind guess

3. Explain breifly what you think the figure discrimination experiment was about and how you think we wanted you to reach the goal of the experiment.

4. In the figure discrimination experiment how sure were you on your correct choices?

a. Absolutely sure
b. Pretty sure
c. Better than 50% sure
d. 50/50 sure
e. Guessing, but kind of sure
f. Guessing
g. Blind guess

5. Do you feel that the tone discrimination experiment helped you perform the figure discrimination experiment in any way? Yes_____, No_____. If so, please explain.

6. Do you think the tone experiment was about something other than what you were told? Yes_____, No_____. Please explain.

7. Do you think the experimenter was trying to get you to do, say, or think something in particular during the tone experiment? Yes_____, No_____. Please explain.

8. Did you try to do, say, or think in the way you feel the experimenter expected during the tone discrimination experiment or did you disregard what you felt he expected?

9. Do you think the experimenter was trying to get you to do, say or think something in particular during the figure discrimination? Yes_____, No_____. If so, what?

10. Did you try to do, say, or think in the way you feel the experimenter expected during the figure discrimination task or did you disregard what you felt he expected?

11. Do you think the figure discrimination task was about something other than what you were told? Yes_____, No_____. What?

12. Which task was the most difficult for you?
a. the tone discrimination
b. the figure discrimination
Why?

13. We are interested in learning about the different learning strategies students use to solve relatively difficult discrimination tasks. How did you go about (what cues did you use, etc.) selecting your answers for:
a. the tone task
b. the figure task

14. Now that the task is done, in what way could you have increased your correct responses?
Please explain.

SET B: Control Subjects

1. Explain briefly what you think the figure discrimination experiment was about and how you think we wanted you to reach the goal of the experiment.

2. In the figure discrimination experiment how sure were you on your correct choices?
a. Absolutely sure
b. Pretty sure
c. Better than 50% sure
d. 50/50 sure
e. Guessing, but kind of sure
f. Guessing
g. Blind guess

3. Did you think the experimenter was trying to get you to do, say, or think something in particular during the figure discrimination? Yes_____, No_____. If so, what?

4. Did you try to do, say, or think in the way you feel the experimenter expected during the figure discrimination task or did you disregard what you felt he expected?

5. Do you think the figure discrimination experiment was about something other than what you were told? Yes_____, No_____. What?

References

Boone, D.C. & McGinley, H. Project LB, an unpublished study. 1975, University of Wyoming.

Brewer, W.F. There is no convincing evidence for operant or classical conditioning in adults humans. In W.B. Weimer and D.S. Palermo (Eds.), Cognition and symbolic processes. Potomac, Md.: Erlbaum, 1974, 1-42.

Bryne, D. The Attraction Paradigm. New York: Academic Press, 1971.

Gerstein, A.I. The effect of reinforcement schedules on meaning generalization and on awareness of the purpose of the experiment. Journal of Personality, 1961, 29, 350-362.

Hare, R.D. Cognitive factors in transfer of meaning. Psychological Reports, 1964, 15, 199-206.

Hare, R.D. Replication report: Cognitive factors in transfer of meaning. Psychological Reports, 1965, 17, 590.

Insko, C.A. & Oakes, W.F. Awareness and the "conditioning" of attitudes. Journal of Personality and Social Psychology, 1966, 4, 487-496.

Lamberth, J. & Byrne, D. Similarity-attraction or demand characteristics? Personality: An International Journal, 1971, 2, 77-91.

McGinley, H. & Boone, D.C. Project DB, an unpublished study. 1974, University of Wyoming.

McGinley, H. & Layton, B. Conditioning of evaluative meaning without contingency or demand awareness. Journal of General Psychology, 1973, 89, 81-90.

McGinley, H. & Layton, B. Affect conditioning in a PA learning list. Paper presented to meeting of Psychonomic Society, San Antonio, Texas, November 1970.

McGinley, H., Layton, B. & McGinley, P. Demand awareness and the conditioning of meaning. Unpublished paper presented to meeting of Rocky Mountain Psychological Association, Denver, Colorado, May 1971.

McGinley, H. & Lloyd, A. Demand awareness and the conditioning of evaluative meaning. Paper presented to meeting of Psychonomic Society, St. Louis, Missouri, November 1971.

Murray, R.P. & McGinley, H. Looking as a measure of attraction. Journal of Applied Social Psychology, 1972, 2, 267-274.

O'Donnell, J.M. The role of awareness in verbal classical conditioning. Paper presented at the Midwestern Psychological Association, Chicago: 1975.

O'Donnell, J.M. & Brown, M.J.K. The classical conditioning of attitudes: A comparative study of ages 8 to 18. Journal of Personality and Social Psychology, 1973, 26, 379-385.

Osgood, C.E., Suci, G.J. & Tannebaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.

Page, M.M. A demand characteristics explanation of operant and classical verbal conditioning. Paper presented at the 83rd Annual Convention of the American Psychological Association, Chicago, 1975. Symposium: Conditioning of Verbal Behavior and Attitudes: A controversy.

Rhine, R.J. A concept-formation approach to attitude acquisition. Psychological Review, 1958, 65, 362-370.

Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton-Century-Crofts, 1966.

Yavuz, H.S. & Bousfield, W.A. Recall of connotative meaning. Psychological Reports, 1959, 5, 319-320.

Zanna, M.P., Kiesler, C.A. & Pilkonis, P.A. Positive and negative attitudinal affect established by classical conditioning. Journal of Personality and Social Psychology, 1970, 14, 321-328.